Homework # 10

1. For $\alpha > 2$ and $0 < a \leq 1$, solve

$$\min_{x\in\mathbb{R}^p} \sum_{i=1}^{p} x_i^\alpha \tag{1}$$
$$\text{subject to } \|x\|^2 = 1,$$
$$x_i \geq a \text{ for } i = 1, 2, \ldots, p.$$

2. (This is a well known problem in econometrics.) Let x_t be the units of cake we eat on day t for $t = 1, \ldots, T$. We require $x_t \geq 0$ for each t and $x_1 + x_1 + \cdots + x_T = 1$. (We always eat a non-negative amount of cake and we have a total of 1 unit of cake to eat.) Define,

$$f(x) = \sum_{t=1}^{T} \beta^{t-1} u(x_t), \tag{2}$$

where $\beta \in (0, 1)$ and $u(w)$ is the amount of utility we derive from eating w units of cake. $f(x)$ models the total utility we derive from eating the cake. Use the KKT conditions to find x that maximizes utility for $u(w)$ given by,

(a) $u(w) = \sqrt{w}$,

(b) $u(w) = w^2$.

(BE CAREFUL, the goal is to MAXIMIZE the utility.)

3. Let $y \in \mathbb{R}^n$ and A an $n \times p$ matrix. Consider the constrained optimization

$$\min_{x\in\mathbb{R}^p} \|y - Ax\|^2 \tag{3}$$
$$\text{subject to } \sum_{i=1}^{p} x_i^2 \leq a^2$$

(a) Use the KKT conditions to solve this optimization.

(b) Consider the ridge regression,

$$\min_{x\in\mathbb{R}^p} \|y - Ax\|^2 + \rho\|x\|^2, \tag{4}$$

for $\rho > 0$. Let x^* be the solution. Show $x^* = (A^T A + \rho I)^{-1} A^T y$. Be sure to show that $(A^T A + \rho I)^{-1}$ always exists.

(c) Assume $(A^T A)^{-1}$ exists. Show that for any $\rho > 0$, if $a = \|x^*\|$ then the constrained optimization is also solved by x^* with Lagrange multiplier equal to ρ. (This provides a $1-1$ correspondence between ridge regression and the constrained optimization in (a). The assumption of $(A^T A)^{-1}$ is not needed and we'll remove it in subsequent work.)

4. Consider the constrained optimization,

$$\begin{aligned} &\min_{x \in \mathbb{R}^p} \|y - Ax\|^2 \\ &\text{subject to } \; x_i \geq 0 \text{ for } i = 1, 2, \ldots, p \end{aligned} \tag{5}$$

Write code that implements a projected gradient approach for solving this problem. Construct and test your code for various y, A combinations. (Don't forget to monitor the loss function!)